Supplementary Materials for

Genetic interaction mapping informs integrative structure determination of protein complexes

Hannes Braberg[#], Ignacia Echeverria[#], Stefan Bohn[#], Peter Cimermancic[#], Anthony Shiver, Richard Alexander, Jiewei Xu, Michael Shales, Raghuvar Dronamraju, Shuangying Jiang, Gajendradhar Dwivedi, Derek Bogdanoff, Kaitlin K. Chaung, Ruth Hüttenhain, Shuyi Wang, David Mavor, Riccardo Pellarin, Dina Schneidman, Joel S. Bader, James S. Fraser, John Morris, James E. Haber, Brian D. Strahl, Carol A. Gross, Junbiao Dai, Jef D. Boeke*, Andrej Sali*, Nevan J. Krogan*

Correspondence to: nevan.krogan@ucsf.edu, sali@salilab.org and
jef.boeke@nyulangone.org

This PDF file includes:

> Figs. S1 to S11
> Tables S3, S8 and S9
> References 1-144 (Main and Supplementary)

Other Supplementary Materials for this manuscript include the following:

> Table S1. Strains used in this study
> Table S2. RNA-seq profiled strains
> Table S4. Genetic interactions of specific modifier-residue pairs
> Table S5. Gene process annotations and FDR values
> Table S6. Spontaneous mutation frequencies and H3K56 acetylation levels
> Table S7. *STE11* 5' and 3' transcript abundance qPCR data
> Data S1. Clustered histone pE-MAP and pairwise S-scores
> Data S2. Nucleosome structure including hypothetical tail structure, modified
> from PDB 1ID3
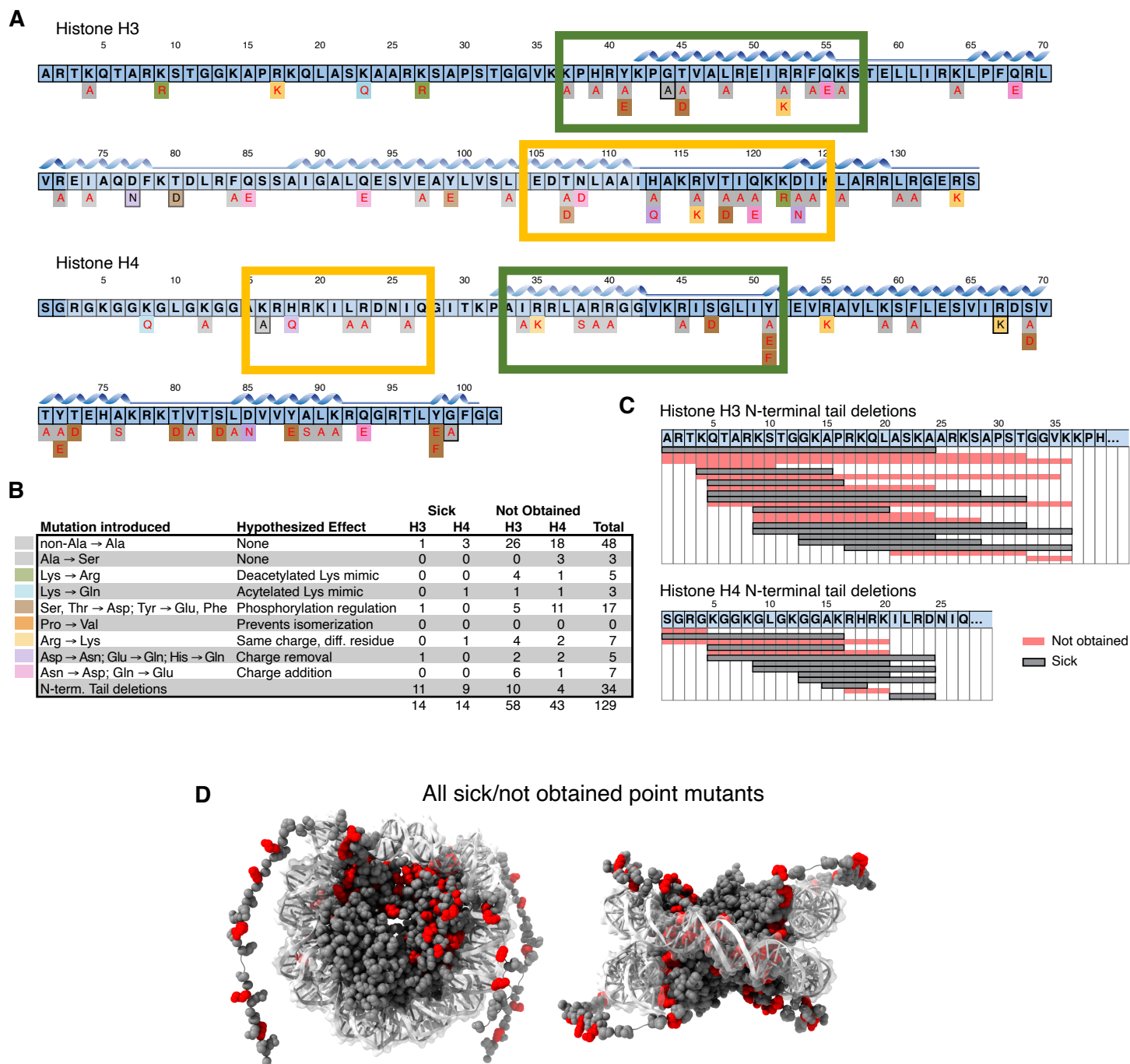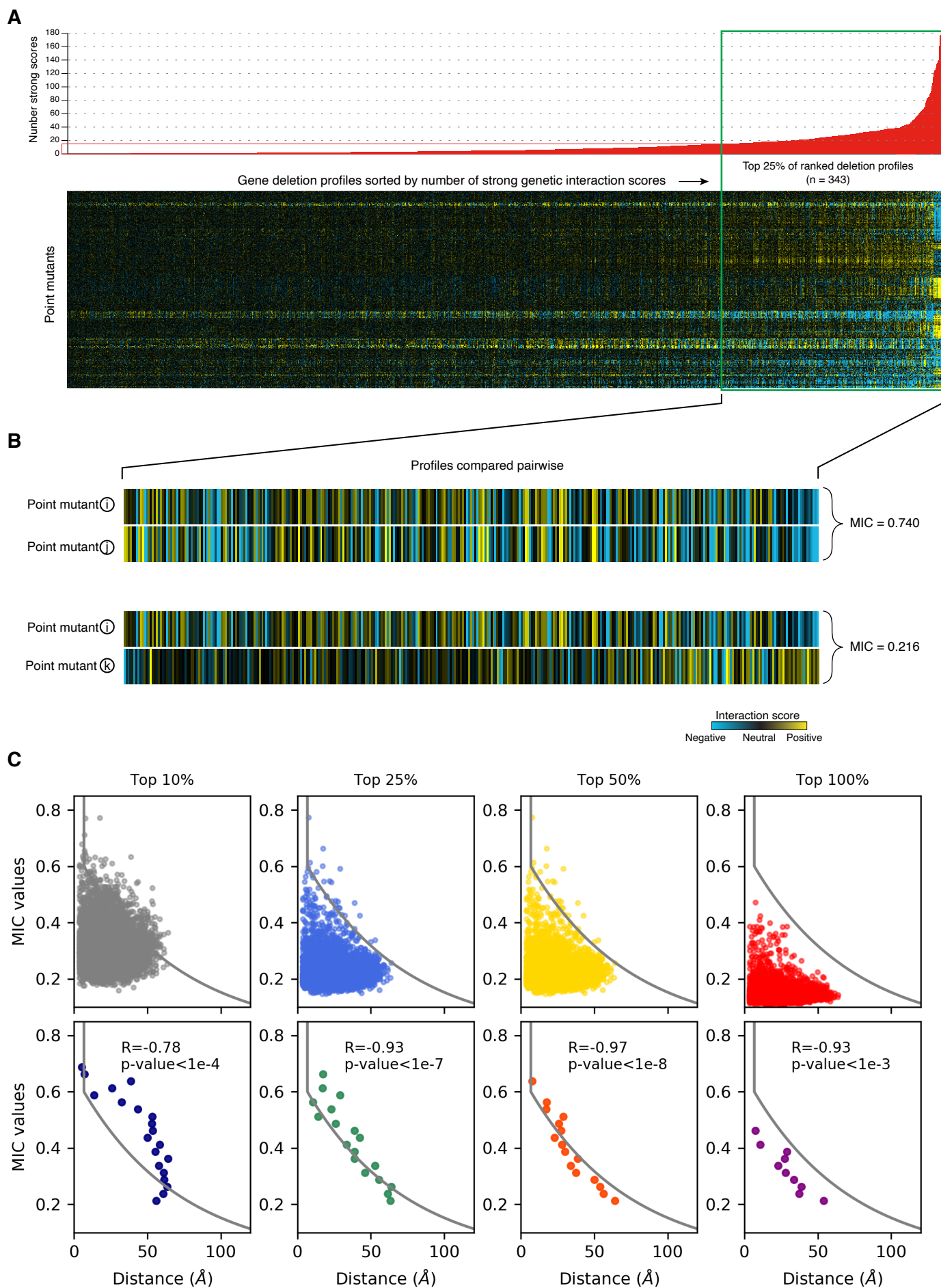> Data S3. Clustered correlation map and pairwise Pearson correlation coefficients

Fig. S1

Fig. S1. Sick and non-viable histone mutant alleles. **(A)** Schematic of the histone point mutants analyzed in this study that were either lethal, could not be constructed after two attempts, or that could not be screened in the pE-MAP (e.g. due to slow growth). Secondary structure elements are indicated as ribbons above the amino acid sequence. The mutation background highlights are color-coded according to the mutation introduced (as in **Fig. S1B**), and the mutation font color indicates whether the mutant was lethal (red) or sick (black). Areas with a high incidence of sick and lethal mutations are highlighted (green boxes: nucleosome entry site; yellow boxes: close to dyad axis). **(B)** Table of sick/lethal histone mutants and their hypothesized effects (color-coding as in **Fig. S1A**). **(C)** Overview of H3 and H4 tail deletion mutants in this study that were either lethal (red) or sick (gray). **(D)** Structural mapping of sick and lethal point mutants (PDB: 1ID3 with hypothetical N-terminal H3 and H4 tails, **Data S2**). Lethal or sick alleles not included in the pE-MAP are highlighted in red.
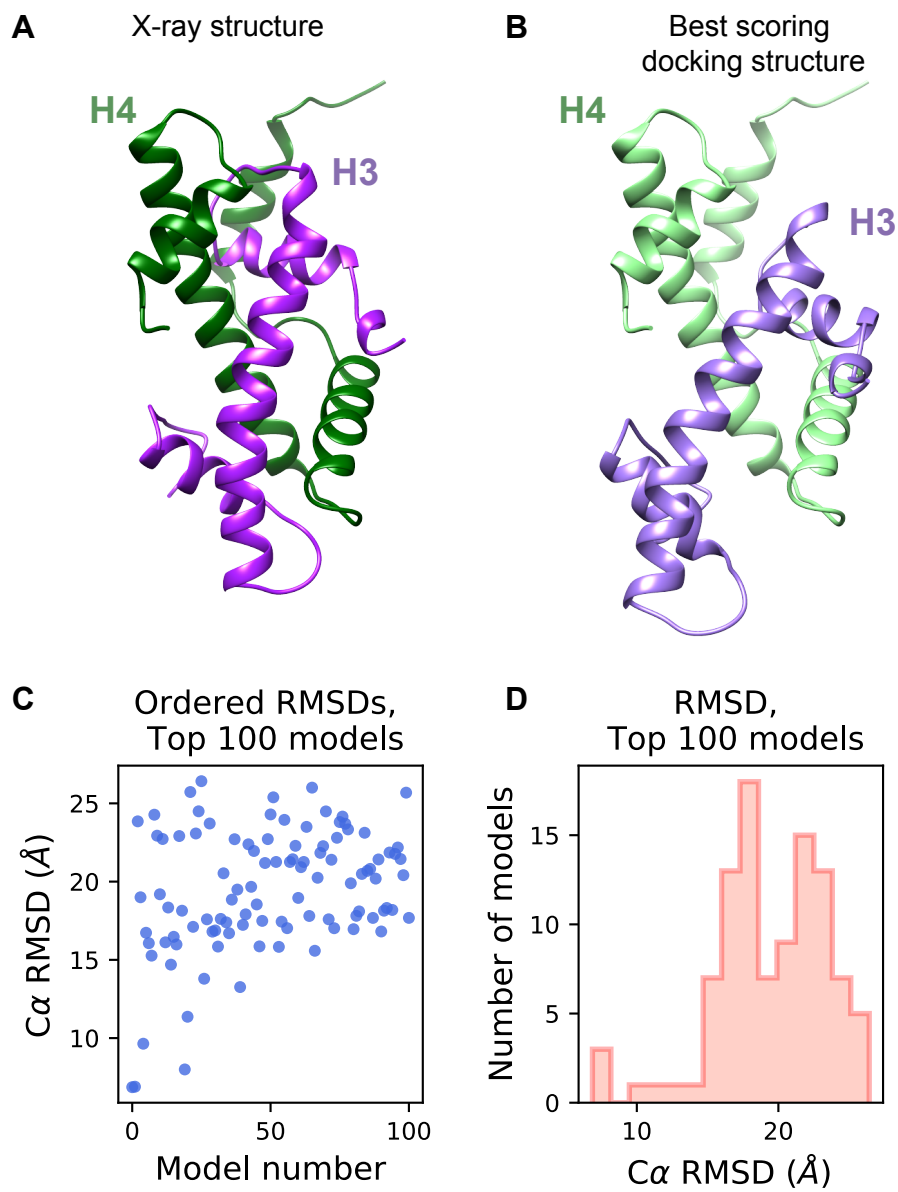
A

B

Signal of histone mutants (sum of absolute S-scores)
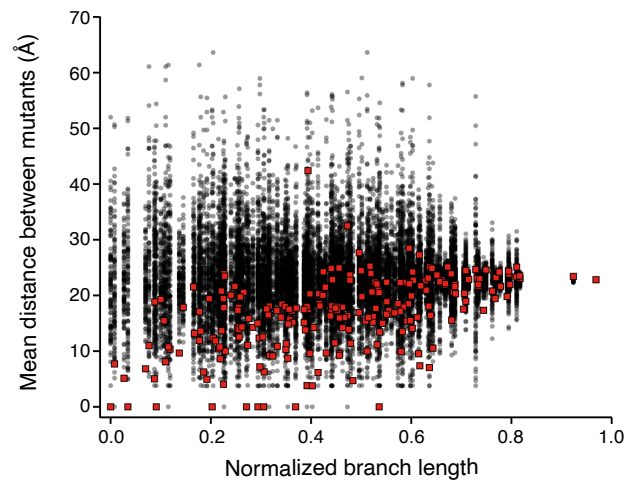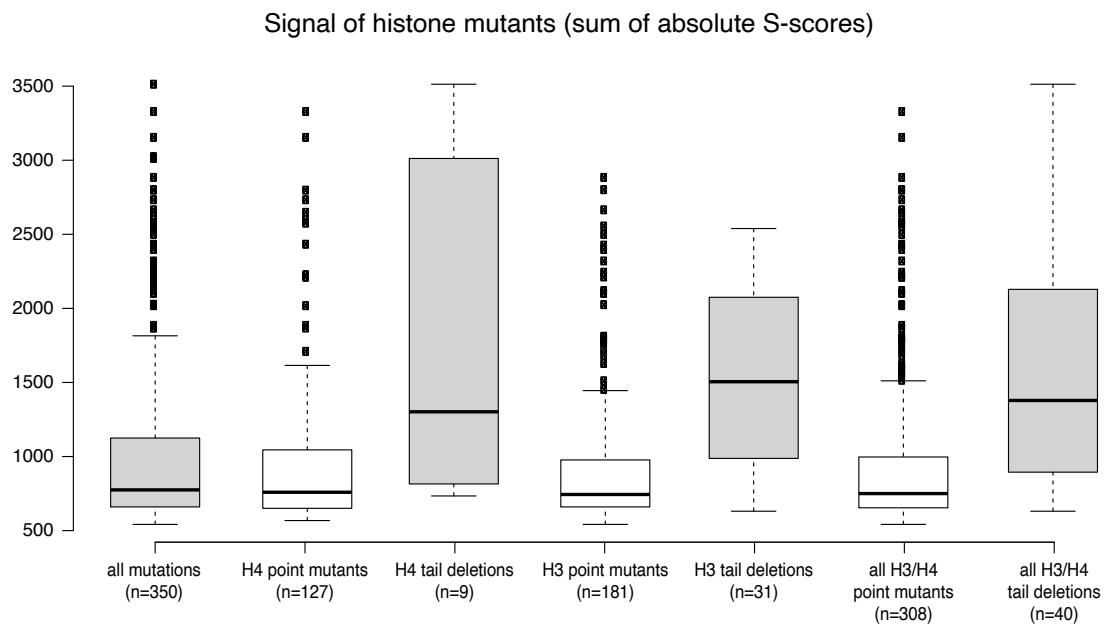
Fig. S2

Fig. S2. Structural trends of the histone pE-MAP. **(A)** The mean distance between histone mutants belonging to a cluster node of the hierarchically clustered pE-MAP plotted against normalized branch length (red; reference nucleosome PDB: 1ID3). The pE-MAP data was filtered prior to clustering to only contain alleles mapping to residues included in 1ID3. The mean distance between mutants for 100 randomly generated trees is plotted in black. **(B)** Comparison of the signal of several histone mutant categories as determined in this dataset. Here, signal is defined as the sum of absolute S-scores. The point mutant boxplots only include single point mutations, and thus exclude two strains with multiple mutations (H3 K4,9,14,18Q and H3 K4,9,14,18A).

A

Gene deletion profiles sorted by number of strong genetic interaction scores →

Top 25% of ranked deletion profiles (n = 343)

B

Profiles compared pairwise

Point mutant (i)
Point mutant (j)
MIC = 0.740

Point mutant (i)
Point mutant (k)
MIC = 0.216

Interaction score
Negative   Neutral   Positive

C

Top 10%   Top 25%   Top 50%   Top 100%

R=-0.78
p-value<1e-4

R=-0.93
p-value<1e-7

R=-0.97
p-value<1e-8

R=-0.93
p-value<1e-3

MIC values

Distance (Å)

Fig. S3

Fig. S3. Processing of pE-MAP data and calculation of phenotypic similarities. (A) To increase the signal-to-noise ratio of the pE-MAP data, the gene deletion profiles were ranked based on the counts of their genetic interaction scores that fell in either the top 2.5% of positive scores or bottom 5% of negative scores of the complete pE-MAP. Gene deletions with the same count were then ranked by the mean of the absolute values of their highest and lowest score. The top 25% of the ranked deletions were retained for computing the point mutant phenotypic profile similarities (B). **(B)** Comparison of phenotypic profiles using MIC. Pairs of point mutants with more similar profiles (*i* and *j*) receive a higher MIC value than those that have less similar profiles (*i* and *k*). **(C)** Statistical association of the distance between two mutated residues with their phenotypic similarity. Top: MIC values were computed after ranking gene deletions and selecting top 10, 25, 50, and 100% of the ranked deletions, and plotted against the distance between the two mutated residues in the x-ray structure (PDB: 1ID3). The grey lines correspond to the upper distance bound used for the implemented distance restraint (**Methods**, Eq. 1). Bottom: Maximum distance for each of the 20 MIC bins. R and p-values correspond to the Pearson correlation coefficient and association significance, respectively, for the log-transformed MIC values.

A X-ray structure

B Best scoring docking structure

C Ordered RMSDs, Top 100 models

D RMSD, Top 100 models

H4 H3

Cα RMSD ($\mathring{A}$)

Model number

Number of models

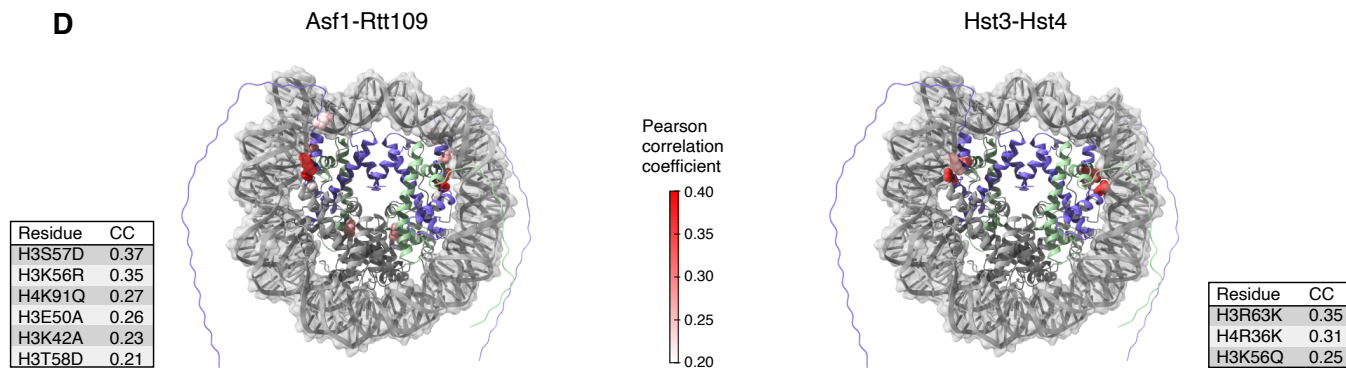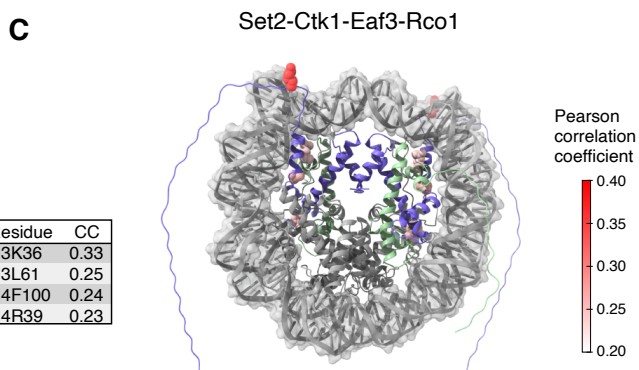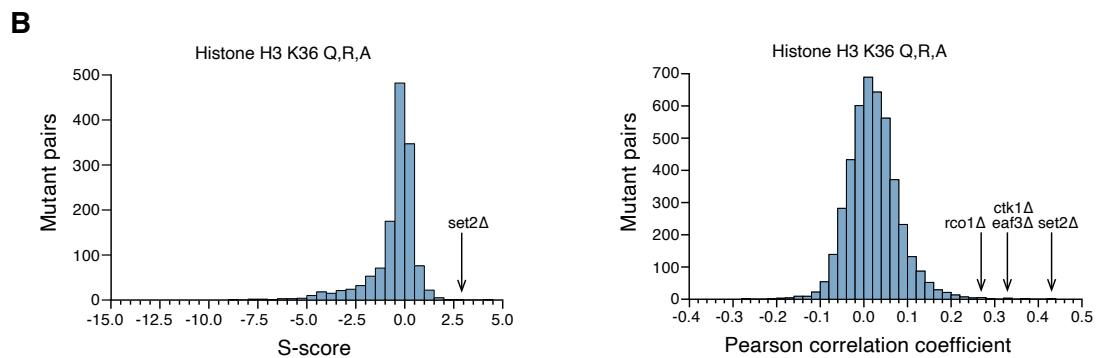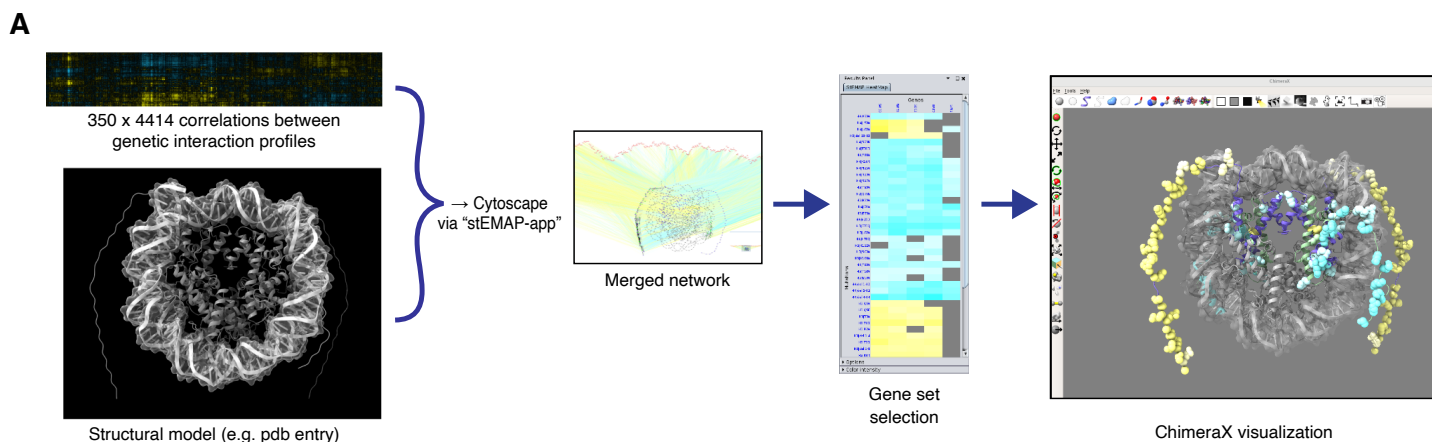Cα RMSD ($\mathring{A}$)

Fig. S4

Fig. S4. Histone H3-H4 docking results. (A) X-ray structure of the histone H3-H4 dimer (PDB: 1ID3). **(B)** Best scoring structure computed by PatchDock. **(C)** Cα RMSD for the top 100 scoring structures computed by PatchDock. **(D)** Cα RMSD histogram for these structures.
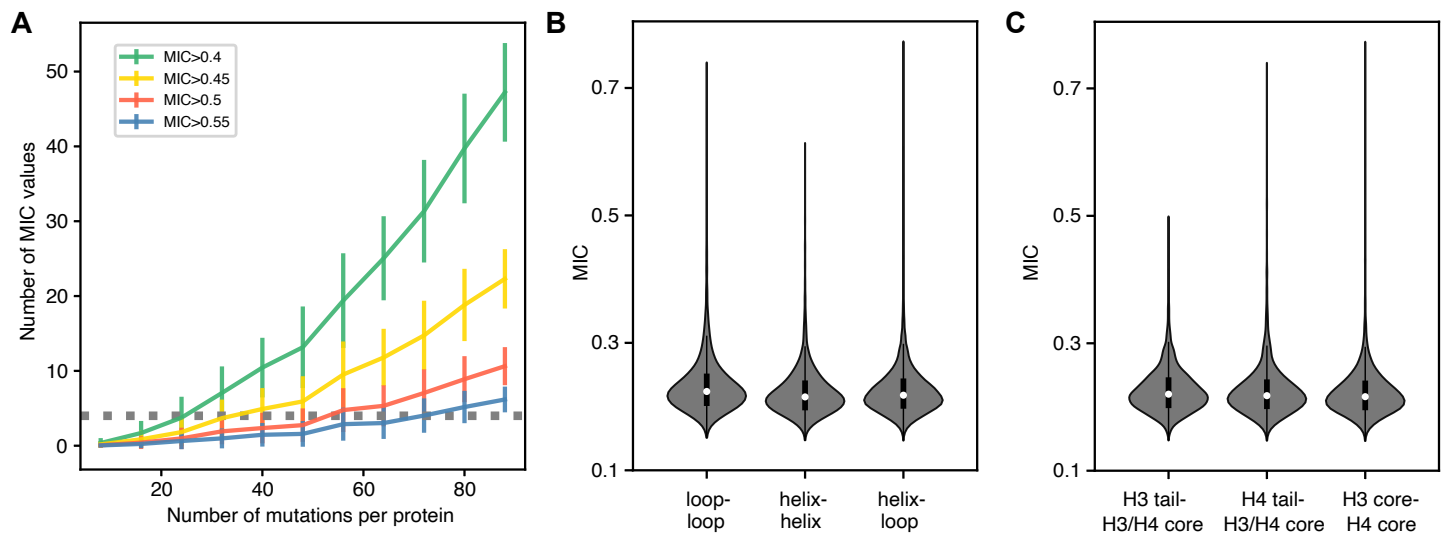
Fig. S5

Fig. S5. Histone MIC value distributions. (A) Relationship between the number of protein-wide systematic mutations (often to alanine) in the H3 and H4 subunits and the number of MIC values above the selected MIC value thresholds. Error bars represent one standard deviation from different random sets of mutations in each subunit. The horizontal gray dashed line represents 4 MIC values above the MIC thresholds. **(B)** Violin plot showing the MIC value distributions when grouped based on the secondary structure of the residue pairs. **(C)** Violin plot showing the MIC value distributions when grouped based on residue pairs being part of the histone cores or tails.

A

350 x 4414 correlations between genetic interaction profiles

Structural model (e.g. pdb entry)

→ Cytoscape via "stEMAP-app"

Merged network

Gene set selection

ChimeraX visualization

B

Histone H3 K36 Q,R,A

set2Δ

Histone H3 K36 Q,R,A

rco1Δ ctk1Δ eaf3Δ set2Δ

C Set2-Ctk1-Eaf3-Rco1

Residue	CC
H3K36	0.33
H3L61	0.25
H4F100	0.24
H4R39	0.23

Pearson correlation coefficient

0.40
0.35
0.30
0.25
0.20

D Asf1-Rtt109

Residue	CC
H3S57D	0.37
H3K56R	0.35
H4K91Q	0.27
H3E50A	0.26
H3K42A	0.23
H3T58D	0.21

Pearson correlation coefficient

0.40
0.35
0.30
0.25
0.20

Hst3-Hst4

Residue	CC
H3R63K	0.35
H4R36K	0.31
H3K56Q	0.25

Fig. S6

Fig. S6. Structural mapping of residue-specific genetic interaction data. (A) Schematic of the structural mapping of genetic interactions using the Cytoscape stEMAP app. Genetic interaction data (here correlations of genetic interaction profiles) and the structure file (modified PDB 1ID3, **Data S2**) are imported into Cytoscape, creating a residue interaction network (RIN) of the nucleosome and the genes it interacts with. The RIN is constructed to reflect the orientation of the 3D view in ChimeraX and the genes are organized to reflect the dendrogram in the original clustered heatmap. The edges are colored to show the interaction (or correlation) scores. Using setsApp, known complexes are loaded to provide a quick way to select the genes in a complex and reflect those genes in all windows. The stEMAPP app will display a heatmap showing only the selected genes and their interacting residues and allows the user to adjust various parameters via the controls at the bottom of the panel, including the minimum number of interactions required. The structural view is shown in the ChimeraX window. When a gene or set of genes are selected, the residues that pass the user defined genetic interaction threshold are shown as spheres and colored according to the interaction score (**Methods**). Selection in the ChimeraX window is linked to the RIN shown in the main window. **(B)** Average distributions of genetic interaction scores (left) and genetic interaction profile correlations (right) of H3K36 mutants (mean of H3K36A, H3K36R, H3K36Q). Genes required for *SET2*-mediated H3K36-methylation that exhibit a mean S-score >2.5 or a mean correlation >0.2 with H3K36 mutants are highlighted. **(C)** Mapping of genetic interaction profile correlations to *SET2* and associated genes, required for H3K36 methylation, on the structure of the nucleosome (modified PDB 1ID3, **Data S2**). N-terminal tail residues of H3 and H4 not included in 1ID3 are visualized as strings on the periphery. Only residues that exhibit a median genetic profile correlation >0.2 with deletions of the *SET2* gene set members (*SET2, CTK1, EAF3* and *RCO1*) are highlighted (**Methods**). H3 (purple), H4 (light green), H2A/H2B and DNA (grey). The red color gradient reflects the strength of the correlation between each residue and the *SET2* gene set, calculated as the median correlation between the residue's tested mutations and the deletions of the *SET2* gene set members. **(D)** Mapping of genetic interaction profile correlations of genes required for H3K56 acetylation and deacetylation on the structure of the nucleosome (modified PDB 1ID3, **Data S2**). N-terminal tail residues of H3 and H4 not included in 1ID3 are visualized as strings on the periphery. Only alleles that exhibit a median genetic interaction profile correlation >0.2 with deletions of *RTT109/ASF1* (responsible for H3K56 acetylation, left), or *HST3/HST4* (responsible for H3K56 deacetylation, right) are highlighted (**Methods**). H3 (purple), H4 (light green), H2A/H2B and DNA (grey). The red color gradient reflects the strength of the correlation between each allele and deletions of *RTT109/ASF1* or *HST3/HST4*, calculated as the median correlation between the allele and the respective gene sets.
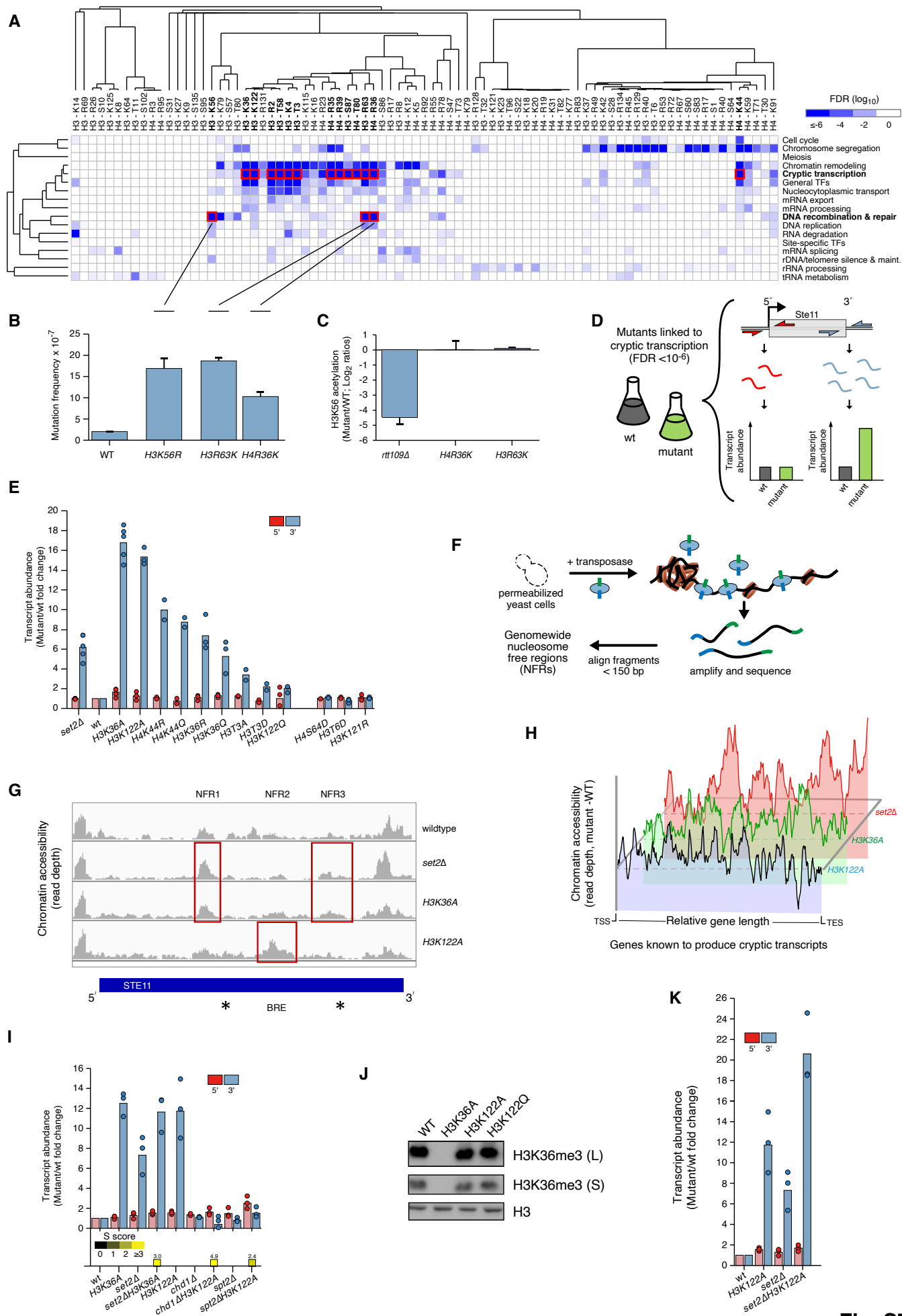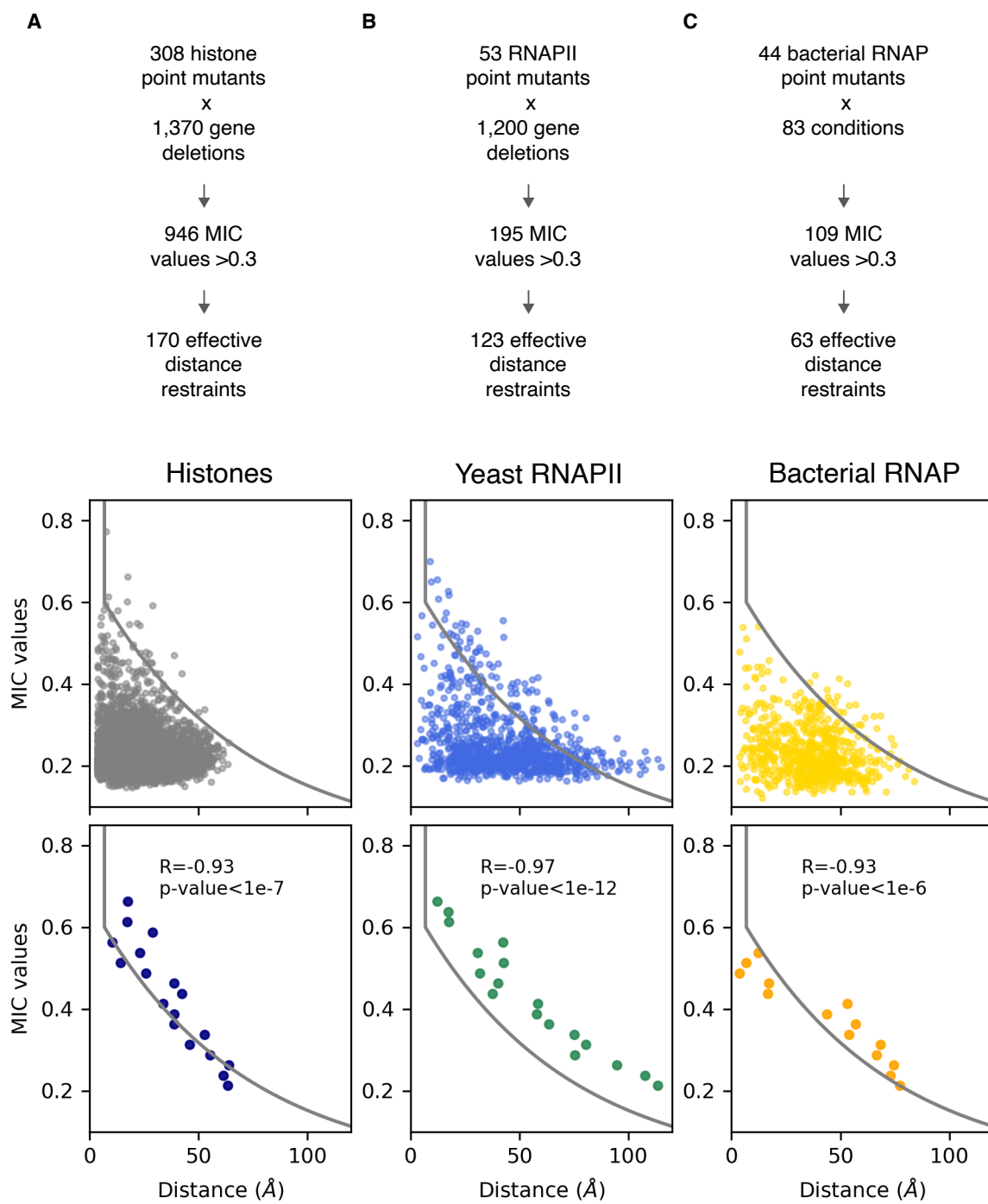
Fig. S7

Fig. S7. Genetic interactions connect H3 and H4 residues to cellular pathways. **(A)** Heatmap representation of gene set enrichment analysis (GSEA, **Methods**) performed to unveil functional connections between histone residues and biological processes related to nuclear function. The genetic interaction profiles of 350 histone alleles were correlated to 4414 genetic interaction profiles from previous studies. The resulting matrix of 350x4414 Pearson correlation coefficients (CCs) was used for GSEA. Only modifiable histone residues are included in the heatmap. The color indicates the False Discovery Rate (FDR) between each residue and biological process. Residues, processes and connections that are further detailed in **Fig. S7B-K** are highlighted in red. **(B)** Residues connected to DNA recombination & repair at FDR < 10^{-6} were ranked by mean correlation, and K->R and R->K mutations of the top 3 residues were examined for their effect on mutation frequency. Spontaneous mutation frequency was measured by a 5-FOA resistance assay at the *URA3* locus (**Methods**). Both H3R63K and H4R36K exhibit strongly increased mutation frequencies at the *URA3* locus compared to wt, at a comparable level to H3K56R. Bar heights represent the mean of three replicates and error bars indicate standard error of the mean (SEM). **(C)** Effect on K56ac-levels of H4R36K and H3R63K, measured by quantitative mass spectrometry. This assay tests if the roles of H4R36 and H3R63 in DNA repair are related to H3K56 acetylation (which has a central role in DNA repair). The lack of effect indicates that H3R63 and H4R36 function via a mechanism independent of H3K56 acetylation. Rtt109 is required for H3K56 acetylation and its deletion serves as a positive control (*141*). Bar heights represent the mean of three replicates and error bars indicate SEM. **(D)** Workflow for cryptic transcription assay. Mutants predicted by GSEA to be involved in cryptic transcription (FDR < 10^{-6}) were assayed for 5'- and 3'-transcript abundance by qPCR at the *STE11* gene, known to produce cryptic transcripts (*33, 142*). A greater output of transcripts from the 3' region than the 5' region indicates the presence of transcription start sites within the gene that give rise to cryptic transcripts. **(E)** 5' (red) and 3' (blue) transcript abundance changes in histone mutants compared to wt (fold-change) at the *STE11* gene. *set2Δ* is shown as a positive control, followed by the mutants predicted to exhibit cryptic transcription (FDR < 10^{-6}) that result in over two-fold change of 3' transcript abundance (H3K36A to H3K122Q), and finally three mutants not predicted to exhibit cryptic transcription (H4S64D, H3T6D and H3K121R). Circles represent replicates and bar heights indicate the geometric means. **(F)** ATAC-seq workflow. ATAC-seq is used to determine nucleosome-free regions (NFRs) on a genome-wide scale. H3K122 is at the histone-DNA interface (<5 Å), and its acetylation by Brd4 leads to nucleosome eviction in human cells (*143*). Nucleosome eviction can be facilitated by histone PTMs or amino acid substitutions that destabilize DNA-histone interactions (*144*). To determine the mechanism of cryptic transcription in H3K122A we used ATAC-seq to map nucleosome free regions (NFRs) in the gene body of *STE11* and other genes known to give rise to cryptic transcripts. **(G)** ATAC-seq reads from open chromatin regions in the gene body of *STE11* (blue). Deletion of *SET2* or a H3K36A mutation give rise to open chromatin at two NFRs, located near TATA-box like sequences (NFR1 and NFR3). H3K122A gives rise to a third NFR in between the other two (NFR2), located at a B-recognition element (BRE). NFRs overlapping with previously reported start sites of cryptic transcripts and sequences important for transcription initiation are highlighted (NFR - red boxes, * - TATA-box element, BRE – B response element). **(H)** Gene body plots of open chromatin regions within 11 genes (*FLO8, AVO1, LCB5, SMC3, SPB4, APM2, DDC1, SYF1, OMS1, PUS4, STE11*) known

to give rise to cryptic transcripts. The plots indicate the difference between each mutant and wt and the dashed lines at 0 indicate the wt chromatin accessibility. The gene lengths were normalized by binning all genes to a set number of intervals and plotting the average normalized read count for all the base pairs in each interval (**Methods**). Consistent with our findings at *STE11*, we observe that *set2Δ*, H3K36A and H3K122A exhibit increased chromatin accessibility compared to wt. **(I)** 5' (red) and 3' (blue) transcript abundance changes in mutants compared to wt (fold-change) at the *STE11* gene. Measured S-scores of double mutants are displayed above the labels. Circles represent replicates and bar heights indicate the geometric means. **(J)** H3K36-trimethylation (H3K36me3) levels as determined by Western Blot of H3K36A, H3K122A and H3K122Q compared to wt H3K36me3 levels. H3K122A has no effect on H3K36 methylation, indicating that the H3K122A-Chd1-Spt2 pathway is independent of Set2-K36me. Histone H3 was used as loading control. L: long exposure; S: short exposure. **(K)** 5' (red) and 3' (blue) transcript abundance changes in mutants compared to wt (fold-change) at the *STE11* gene. Circles represent replicates and the bar heights indicate the geometric means. *set2Δ* H3K122A exhibits a much greater 3' transcript abundance change than either single mutant alone, supporting that the H3K122A-Chd1-Spt2 pathway is independent of Set2-K36me.

A

308 histone
point mutants
x
1,370 gene
deletions

↓

946 MIC
values >0.3

↓

170 effective
distance
restraints

B

53 RNAPII
point mutants
x
1,200 gene
deletions

↓

195 MIC
values >0.3

↓

123 effective
distance
restraints

C

44 bacterial RNAP
point mutants
x
83 conditions

↓

109 MIC
values >0.3

↓

63 effective
distance
restraints

Histones

Yeast RNAPII

Bacterial RNAP

R=-0.93
p-value<1e-7

R=-0.97
p-value<1e-12

R=-0.93
p-value<1e-6

MIC values

Distance (Å)

Fig. S8

Fig. S8. Statistical association of the distance between two mutated residues with their phenotypic similarity for different pE-MAP/CG-MAP datasets. (A) Top: Summary of the genetic interaction data used for integrative structure determination for histones H3-H4. Only the 308 single point mutations were included in this analysis. 40 tail deletions and two strains with point mutations in multiple locations (H3 K4,9,14,18Q and H3 K4,9,14,18A) were excluded. Middle: Scatter plot of the MIC values and distances between the mutated residues in the wt histone structure (PDB: 1ID3). Bottom: Plot of the upper distance bound obtained by binning the MIC values into 20 intervals and selecting the maximum distance spanned by any pair of residues in each bin, followed by fitting a logarithmic decay function to these maximum distances (grey line, Eq. 1). The R-values and p value are reported for the Pearson correlation between the distances and the log-transformed MIC values (Eq.1). **(B)** Same as A, for yeast RNAPII. The complete RNAPII pE-MAP contains data from 53 point mutations, but only 44 of these target Rpb1 or Rpb2. The remaining 9 mutations are in other subunits and included in the MIC value count but not used for restraints. Distances were retrieved from PDB 2E2H. The grey line is shown for reference and corresponds to the fit to the histone upper distance bound. **(C)** Same as A, for subunits RpoB and RpoC of bacterial RNAP. Distances were retrieved from PDB 4YG2. The grey line is shown for reference and corresponds to the fit to the histone upper distance bound.
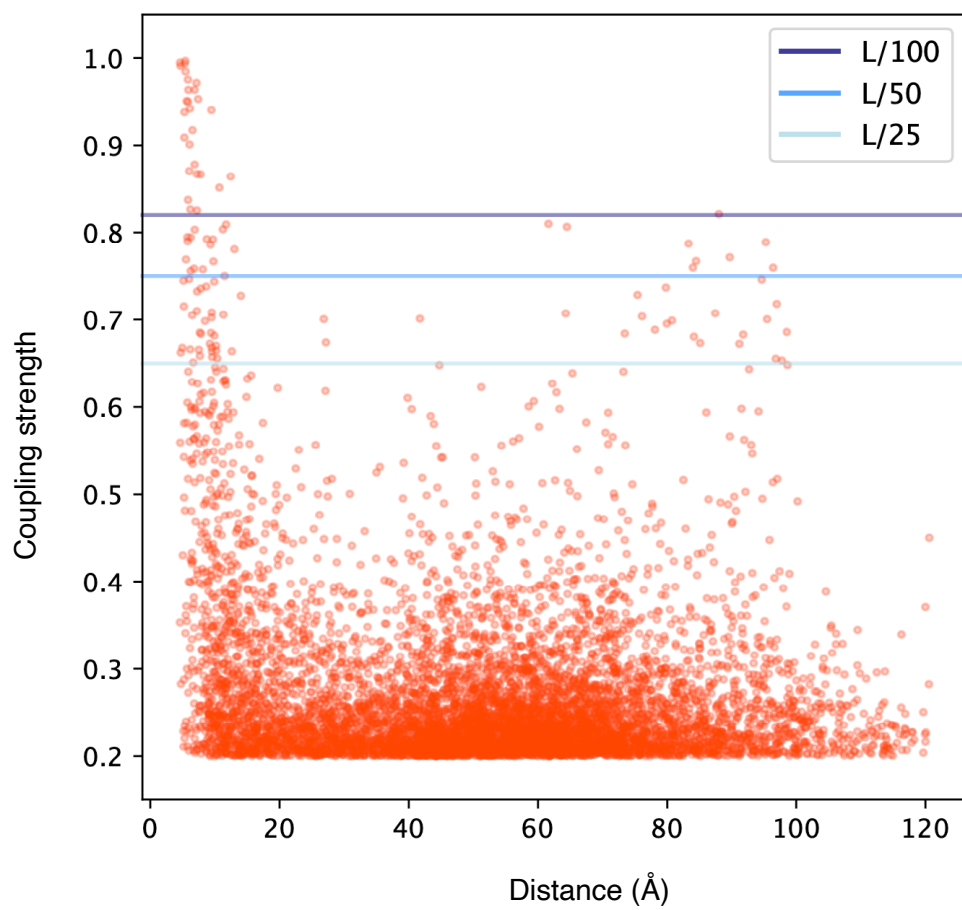
Fig. S9
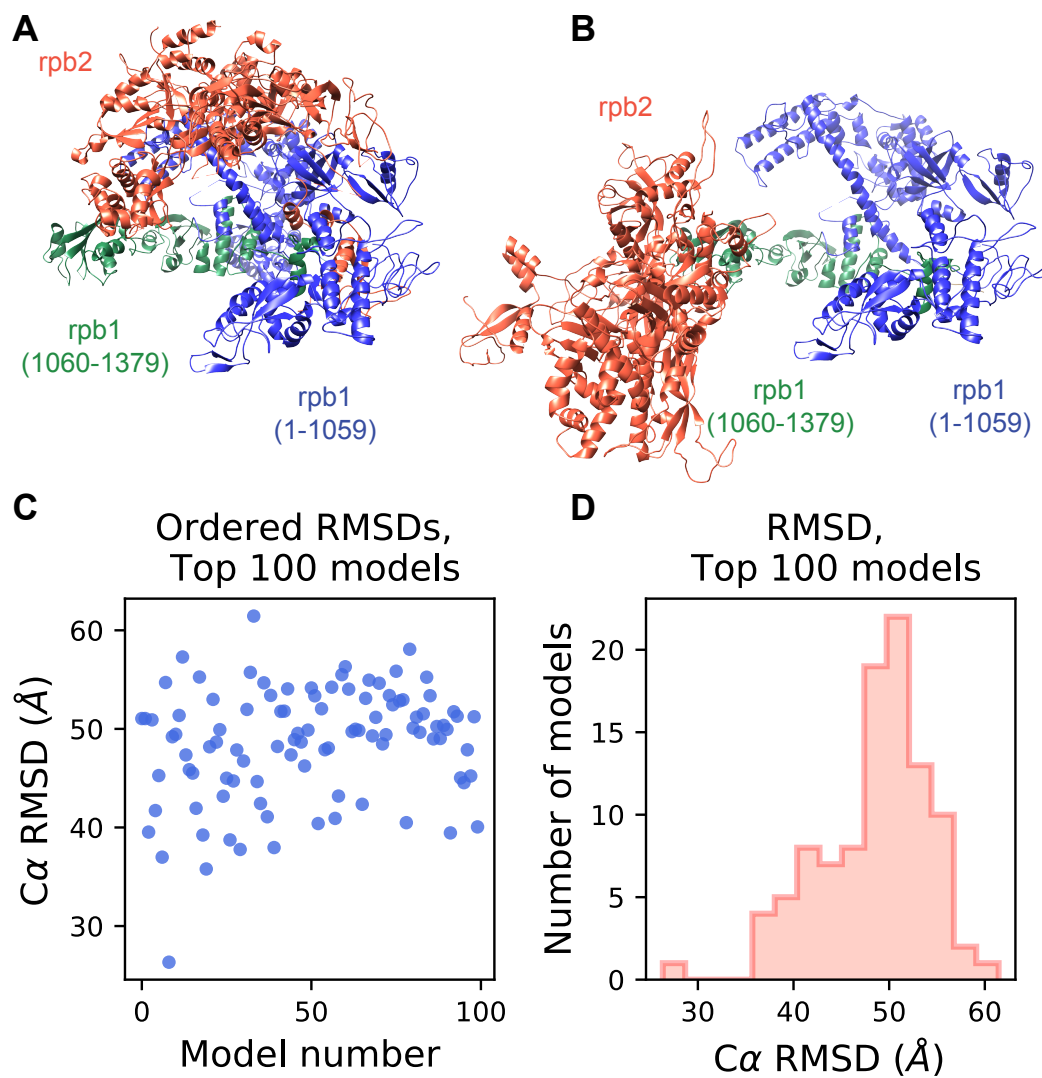
Fig. S9. Coupling strengths for RpoB and RpoC predicted by RaptorX ComplexContact. Dependance of coupling strengths on the distances between residues in the X-ray structure (PDB: 4YG2) is plotted. Horizontal lines correspond to the coupling strength cutoffs when considering the top L/100, L/50, and L/25 predicted contacts, where L is the length of the concatenated sequence.
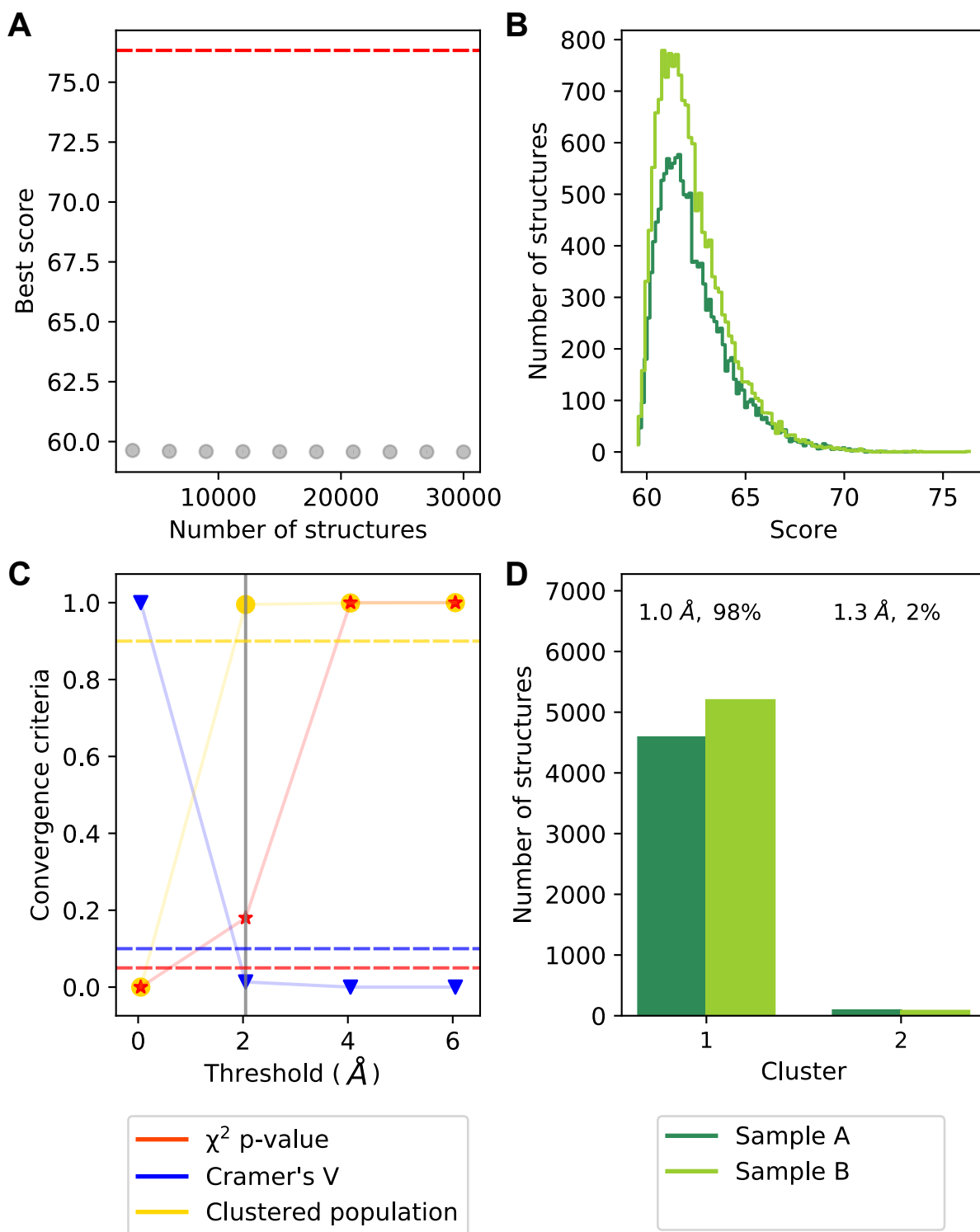
Fig. S10

Fig. S10. Estimation of sampling precision for histones H3-H4. (A) Convergence of the model scores in the ensemble. Grey dots show that the scores do not continue to improve as more structures are independently computed. The dotted line indicates the highest score in the ensemble. **(B)** Distribution of scores for structures in samples A (dark green) and B (light green), comprising 20,000 random models in the ensemble. The magnitude of the difference is small, as demonstrated by the Kolmogorov-Smirnov two-sample test statistic (D=0.13). **(C)** Three criteria for determining the sampling precision (y-axis), evaluated as a function of the RMSD clustering threshold (x-axis). First, the p-value is computed using the χ2-test (one-sided) for homogeneity of proportions (red stars). Second, an effect size for the χ2-test is quantified by the Cramer's *V* value (blue triangles). Third, the population of structures in sufficiently large clusters (containing at least ten structures from each sample) is shown as yellow circles. The vertical dotted grey line indicates the RMSD clustering threshold at which three conditions are satisfied (χ2-test p-value (0.679) > 0.05 (red, horizontal dotted line), Cramer's *V* (0.06) < 0.10 (blue, horizontal dotted line), and the population of clustered structures (0.98) > 0.80 (yellow, horizontal dotted line)), thus defining the sampling precision of 2.03 Å. The three solid curves (in red, blue, and yellow) were drawn through the points to help visualize the results. **(D)** Population of structures in samples A and B in each of the three clusters obtained by threshold-based clustering using an RMSD threshold of 2.05 Å. The dominant cluster (cluster 1) contains 98% of the structures. Cluster precision and population is shown for each cluster. The precision of the dominant cluster defines the model precision.

A

rpb2

rpb1
(1060-1379)

rpb1
(1-1059)

B

rpb2

rpb1
(1060-1379)

rpb1
(1-1059)

C

Ordered RMSDs,
Top 100 models

Cα RMSD (Å)

Model number

D

RMSD,
Top 100 models

Number of models

Cα RMSD (Å)

Fig. S11

Fig. S11. RNAPII docking results. (A) X-ray structure of Rpb1-Rpb2 (PDB: 2E2H). **(B)** Best scoring docking structure computed by PatchDock. **(C)** Cα RMSD for the top 100 scoring models obtained using rigid body docking structures computed by PatchDock. **(D)** Cα RMSD histogram for these top 100 scoring models obtained using rigid body docking.

Table S3: Summary of Integrative Structure Determination of the Histone H3-H4 dimer

1) Gathering information

Prior models	h3: comparative model, template 1TZY:G
	h4: comparative model, template 1TZY:H
Physical principles and statistical preferences	Excluded volume
	Sequence connectivity
Experimental data	170 pE-MAP derived distance restraints

2) Representing the system

Composition (number of copies)	h3: 1
	h4: 1
Atomic (structured) components	h3: 39-136
	h4: 20-103
Unstructured components	None
Resolution of structured components	1 [R1] residue per bead
Resolution of unstructured components	None
Structural coverage	100.0 %
Rigid body (RB) definitions	RB1: h3$_{39-136}$
	RB2: h4$_{20-103}$
Spatial restraints encoded into scoring function	Excluded volume; applied to the R1 representation
	Sequence connectivity; applied to the R1 representation
	pE-MAP MIC pair-restraints; applied to the R1 representation

3) Structural Sampling

Sampling method	Replica Exchange Gibbs sampling, based on Metropolis Monte Carlo
Replica exchange temperature range	1.0 - 2.5
Number of replicas	4
Number of runs	60
Number of structures generated	2500000
Movers for flexible string of bead	Random translation up to 4.0 Å
CPU time	4 hours on 20 processors

4) Validating the model

Models selected for validation

Number of models after equilibration	2500000
Number of models that satisfy the input information	1162507
Number of structures in samples A/B	639091/523416
p-value of non-parametric Kolmogorov-Smirnov two-sample test	0.006 (threshold p-value > 0.05)
Kolmogorov-Smirnov two-sample test statistic, D	0.96

Thoroughness of the structural sampling

Sampling precision	2.03 Å
Homogeneity of proportions χ^2 test (p-value)/Cramers V value	0.679/0.006 (thresholds: p-value>0.05 OR Cramer's V<0.1)
Number of clusters	1
Cluster populations	cluster 1 : 98.6 %
Cluster precisions	cluster 1 : 1.04 Å
Average cross-correlation between localization probability densities of samples A and B	cluster 1: 1.0

Validation by information used for modeling

Percent of sequence connectivity restraints satisfied per structure	99 %
Percent pE-MAP restraints satisfied per structure	87 %
Percent of excluded volume restraints satisfied per structure	99 %

5) Benchmark

Structural accuracy (95 % CI)	3.8 (2.6-5.1) Å
PDB used for benchmark	1ID3

6) Software and data availability

Modeling programs	IMP PMI module, version develop-9c8707cfc5
	Integrative Modeling Platform (IMP), version develop-9c8707cfc5
	MODELLER, version 9.21
Modeling scripts and data	https://integrativemodeling.org/systems/pemap

Homology detection and structure prediction   HHPred, version 2.0.16
Visualization and plotting   UCSF Chimera, version 1.10
                                                       Matplotlib, version 3.0.3

Homology detection and structure prediction   HHPred, version 2.0.16
Visualization and plotting   UCSF Chimera, version 1.10
                                                       Matplotlib, version 3.0.3

Table S3. Summary of the integrative structure determination, thoroughness of configurational sampling, structure precision, and validation of histones H3 and H4. More details in (*112*).

Table S8: Summary of Integrative Structure Determination of RNAPII Rpb1-Rpb2

1) Gathering information	
Prior models	rpb1: comparative model, template 6GMH:A
	rpb2: comparative model, template 4AYB:B
	None: sequence
Physical principles and statistical preferences	Excluded volume
	Sequence connectivity
Experimental data	123 pE-MAP derived distance restraints
2) Representing the system	
Composition (number of copies)	rpb1: 1
	rpb2: 1
Atomic (structured) components	rpb1: 13-58, 82-105, 120-143, 173-184, 200-551, 565-589, 606-1077, 1098-1105, 1119-1171, 1191-1240, 1256-1286, 1304-1404
	rpb2: 24-68, 88-139, 162-301, 314-565, 579-639, 655-661, 681-709, 740-871, 887-917, 934-1099
Unstructured components	rpb1: 1-12, 59-81, 106-119, 144-172, 185-199, 552-564, 590-605, 1078-1097, 1106-1112, 1113-1118, 1172-1190, 1241-1255, 1287-1303
	rpb2: 1-23, 69-87, 140-161, 302-313, 566-578, 640-654, 662-680, 710-739, 872-886, 918-933
Resolution of structured components	1 [R1] residue per bead
Resolution of unstructured components	10 [R10] residues per bead
Structural coverage	85.07 %
Rigid body (RB) definitions	RB1: $rpb1_{1-1105}$
	RB2: $rpb1_{1113-1404}$
	RB3: $rpb2_{1-1099}$
Spatial restraints encoded into scoring function	Excluded volume; applied to the R1 representation
	Sequence connectivity; applied to the R1 representation
	pE-MAP MIC pair-restraints; applied to the R1 representation
3) Structural Sampling	
Sampling method	Replica Exchange Gibbs sampling, based on Metropolis Monte Carlo
Replica exchange temperature range	1.0 - 2.5
Number of replicas	8
Number of runs	60
Number of structures generated	3600000
Movers for flexible string of bead	Random translation up to 4.0 Å
CPU time	36 hours on 20 processors
4) Validating the model	
Models selected for validation	
Number of models after equilibration	3600000
Number of models that satisfy the input information	425380
Number of structures in samples A/B	210367/215013
p-value of non-parametric Kolmogorov-Smirnov two-sample test	0.012 (threshold p-value > 0.05)
Kolmogorov-Smirnov two-sample test statistic, D	0.23
Thoroughness of the structural sampling	
Sampling precision	14.84 Å
Homogeneity of proportions χ^2 test (p-value)/Cramers V value	0.000/0.099 (thresholds: p-value>0.05 OR Cramer's V<0.1)
Number of clusters	2
Cluster populations	cluster 1 : 81.7 %
	cluster 2 : 14.8 %
Cluster precisions	cluster 1 : 9.78 Å
	cluster 2 : 9.92 Å
Average cross-correlation between localization probability densities of samples A and B	cluster 1: 0.76
	cluster 2: 0.84
Validation by information used for modeling	
Percent of sequence connectivity restraints satisfied per structure	99 %
Percent pE-MAP restraints satisfied per structure	95 %
Percent of excluded volume restraints satisfied per structure	99 %

5) Benchmark	
Structural accuracy (95 % CI)	16.8 (12.5-23.2) Å
PDB used for benchmark	1I3Q
6) Software and data availability	
Modeling programs	IMP PMI module, version develop-39c22a602
	Integrative Modeling Platform (IMP), version develop-39c22a602
	MODELLER, version 9.21
Modeling scripts and data	https://integrativemodeling.org/systems/pemap
Homology detection and structure prediction	HHPred, version 2.0.16
Visualization and plotting	UCSF Chimera, version 1.10
	Matplotlib, version 3.0.3

Table S8. Summary of the integrative structure determination, thoroughness of configurational sampling, structure precision, and validation of RNAPII Rpb1-Rpb2. More details in (*112*).

Table S9: Summary of Integrative Structure Determination of Bacterial RNAP subunits Rpob and Rpoc

1) Gathering information	
Prior models	X-ray structure 4YG2
Physical principles and statistical preferences	Excluded volume
	Sequence connectivity
Experimental data	63 conditional genetics derived distance restraints
2) Representing the system	
Composition (number of copies)	rpob: 1
	rpoc: 1
Atomic (structured) components	rpob: 3-1342
	rpoc: 8-931, 1135-1374
Unstructured components	rpob: 1-2
	rpoc: 932-1134
Resolution of structured components	1 [R1] residue per bead
Resolution of unstructured components	20 [R20] residues per bead
Structural coverage	92.49 %
Rigid body (RB) definitions	RB1: rpob$_{3-1342}$
	RB2: rpoc$_{8-931}$,rpoc$_{1135-1374}$
Spatial restraints encoded into scoring function	Excluded volume; applied to the R1 representation
	Sequence connectivity; applied to the R1 representation
	pE-MAP MIC pair-restraints; applied to the R1 representation
3) Structural Sampling	
Sampling method	Replica Exchange Gibbs sampling, based on Metropolis Monte Carlo
Replica exchange temperature range	1.0 - 2.5
Number of replicas	4
Number of runs	50
Number of structures generated	2500000
Movers for flexible string of bead	Random translation up to 4.0 Å
CPU time	16 hours on 20 processors
4) Validating the model	
Models selected for validation	
Number of models after equilibration	2500000
Number of models that satisfy the input information	656275
Number of structures in samples A/B	407275/249000
p-value of non-parametric Kolmogorov-Smirnov two-sample test	0.251 (threshold p-value > 0.05)
Kolmogorov-Smirnov two-sample test statistic, D	0.0
Thoroughness of the structural sampling	
Sampling precision	9.38 Å
Homogeneity of proportions χ^2 test (p-value)/Cramers V value	1.000/0.000 (thresholds: p-value>0.05 OR Cramer's V<0.1)
Number of clusters	1
Cluster populations	cluster 1 : 100.0 %
Cluster precisions	cluster 1 : 6.61 Å
Average cross-correlation between localization probability densities of samples A and B	cluster 1: 0.68
Validation by information used for modeling	
Percent of sequence connectivity restraints satisfied per structure	99 %
Percent pE-MAP restraints satisfied per structure	94 %
Percent of excluded volume restraints satisfied per structure	99 %
5) Benchmark	
Structural accuracy (95 % CI)	15.0 (12.1-18.2) Å
PDB used for benchmark	4YG2
6) Software and data availability	
Modeling programs	IMP PMI module, version develop-39c22a602
	Integrative Modeling Platform (IMP), version develop-39c22a602
Modeling scripts and data	https://integrativemodeling.org/systems/pemap
Homology detection and structure prediction	HHPred, version 2.0.16

Table S9. Summary of the integrative structure determination, thoroughness of configurational sampling, structure precision, and validation of bacterial RNAP RpoB-RpoC. More details in (*112*).

References:

1. F. Alber, F. Forster, D. Korkin, M. Topf, A. Sali, Integrating diverse data for structure determination of macromolecular assemblies. *Annu. Rev. Biochem.* **77**, 443-477 (2008).
2. F. Herzog *et al.*, Structural probing of a protein phosphatase 2A network by chemical cross-linking and mass spectrometry. *Science* **337**, 1348-1352 (2012).
3. M. P. Rout, A. Sali, Principles for Integrative Structural Biology Studies. *Cell* **177**, 1384-1403 (2019).
4. F. Alber *et al.*, Determining the architectures of macromolecular assemblies. *Nature* **450**, 683-694 (2007).
5. D. Russel *et al.*, Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol.* **10**, e1001244 (2012).
6. A. B. Ward, A. Sali, I. A. Wilson, Biochemistry. Integrative structural biology. *Science* **339**, 913-915 (2013).
7. F. Alber *et al.*, The molecular architecture of the nuclear pore complex. *Nature* **450**, 695-701 (2007).
8. K. Lasker *et al.*, Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proc Natl Acad Sci U S A* **109**, 1380-1387 (2012).
9. A. Loquet *et al.*, Atomic model of the type III secretion system needle. *Nature* **486**, 276-279 (2012).
10. Z. Duan *et al.*, A three-dimensional model of the yeast genome. *Nature* **465**, 363-367 (2010).
11. J. M. Plitzko, B. Schuler, P. Selenko, Structural Biology outside the box—inside the cell. *Curr. Opin. Struct. Biol.* **46**, 110-121 (2017).
12. M. Dimura *et al.*, Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems. *Curr. Opin. Struct. Biol.* **40**, 163-185 (2016).
13. S. R. Collins, A. Roguev, N. J. Krogan, Quantitative genetic interaction mapping using the E-MAP approach. *Methods Enzymol.* **470**, 205-231 (2010).
14. P. Beltrao, G. Cagney, N. J. Krogan, Quantitative genetic interactions reveal biological modularity. *Cell* **141**, 739-745 (2010).
15. H. Braberg *et al.*, From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. *Cell* **154**, 775-788 (2013).
16. H. Braberg, E. A. Moehle, M. Shales, C. Guthrie, N. J. Krogan, Genetic interaction analysis of point mutations enables interrogation of gene function at a residue-level resolution: exploring the applications of high-resolution genetic interaction mapping of point mutations. *Bioessays* **36**, 706-713 (2014).
17. N. Halabi, O. Rivoire, S. Leibler, R. Ranganathan, Protein sectors: evolutionary units of three-dimensional structure. *Cell* **138**, 774-786 (2009).
18. D. S. Marks *et al.*, Protein 3D structure computed from evolutionary sequence variation. *PLoS One* **6**, e28766 (2011).
19. G. Diss, B. Lehner, The genetic landscape of a physical interaction. *Elife* **7**, e32472 (2018).
20. A. L. Shiver *et al.*, Chemical-genetic interrogation of RNA polymerase mutants reveals structure-function relationships and physiological tradeoffs. *bioRxiv*, 2020.2006.2016.155770 (2020).
21. H. Huang, S. Lin, B. A. Garcia, Y. Zhao, Quantitative proteomic analysis of histone modifications. *Chemical reviews* **115**, 2376-2418 (2015).
22. J. Dai *et al.*, Probing nucleosome function: a highly versatile library of synthetic histone H3 and H4 mutants. *Cell* **134**, 1066-1078 (2008).

23. S. Jiang *et al.*, Construction of Comprehensive Dosage-Matching Core Histone Mutant Libraries for Saccharomyces cerevisiae. *Genetics* **207**, 1263-1273 (2017).
24. J. E. Brownell *et al.*, Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843-851 (1996).
25. L. K. Durrin, R. K. Mann, P. S. Kayne, M. Grunstein, Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* **65**, 1023-1031 (1991).
26. H. Braberg *et al.*, Quantitative analysis of triple-mutant genetic interactions. *Nature protocols* **9**, 1867-1881 (2014).
27. J. E. Haber *et al.*, Systematic triple-mutant analysis uncovers functional connectivity between pathways involved in chromosome regulation. *Cell Rep* **3**, 2168-2178 (2013).
28. S. R. Collins, M. Schuldiner, N. J. Krogan, J. S. Weissman, A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome biology* **7**, R63 (2006).
29. S. R. Collins *et al.*, Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* **446**, 806-810 (2007).
30. T. Miller *et al.*, COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci U S A* **98**, 12902-12907 (2001).
31. A. Roguev *et al.*, The Saccharomyces cerevisiae Set1 complex includes an Ash2 homologue and methylates histone 3 lysine 4. *EMBO J* **20**, 7137-7148 (2001).
32. G. Mizuguchi *et al.*, ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**, 343-348 (2004).
33. M. J. Carrozza *et al.*, Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581-592 (2005).
34. S. Venkatesh *et al.*, Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. *Nature* **489**, 452-455 (2012).
35. M. C. Keogh *et al.*, Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**, 593-605 (2005).
36. D. N. Reshef *et al.*, Detecting novel associations in large data sets. *Science* **334**, 1518-1524 (2011).
37. D. Albanese *et al.*, minerva and minepy: a C engine for the MINE suite and its R, Python and MATLAB wrappers. *Bioinformatics* **29**, 407-408 (2013).
38. C. L. White, R. K. Suto, K. Luger, Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.* **20**, 5207-5218 (2001).
39. D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H. J. Wolfson, PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363-367 (2005).
40. D. Baker, A. Sali, Protein structure prediction and structural genomics. *Science* **294**, 93-96 (2001).
41. S. J. Kim *et al.*, Integrative structure and functional anatomy of a nuclear pore complex. *Nature* **555**, 475-482 (2018).
42. C. Gutierrez *et al.*, Structural dynamics of the human COP9 signalosome revealed by cross-linking mass spectrometry and integrative modeling. *Proc Natl Acad Sci U S A* **117**, 4088-4098 (2020).
43. K. S. Molnar *et al.*, Cys-scanning disulfide crosslinking and bayesian modeling probe the transmembrane signaling mechanism of the histidine kinase, PhoQ. *Structure* **22**, 1239-1251 (2014).
44. Y. Kwon *et al.*, Structural basis of CD4 downregulation by HIV-1 Nef. *Nat Struct Mol Biol* **27**, 822-828 (2020).

45. C. J. Ryan *et al.*, Hierarchical modularity and the evolution of genetic interactomes across species. *Mol Cell* **46**, 691-704 (2012).
46. J. Wysocka *et al.*, A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* **442**, 86-90 (2006).
47. P. Shannon *et al.*, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).
48. T. D. Goddard *et al.*, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018).
49. D. Schneidman-Duhovny, R. Pellarin, A. Sali, Uncertainty in integrative structural modeling. *Curr Opin Struct Biol* **28**, 96-104 (2014).
50. N. L. Maas, K. M. Miller, L. G. DeFazio, D. P. Toczyski, Cell cycle and checkpoint regulation of histone H3 K56 acetylation by Hst3 and Hst4. *Mol Cell* **23**, 109-119 (2006).
51. J. Han *et al.*, A Cul4 E3 ubiquitin ligase regulates histone hand-off during nucleosome assembly. *Cell* **155**, 817-829 (2013).
52. T. Tsubota *et al.*, Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. *Mol Cell* **25**, 703-712 (2007).
53. E. M. Hyland *et al.*, Insights into the role of histone H3 and histone H4 core modifiable residues in Saccharomyces cerevisiae. *Mol Cell Biol* **25**, 10060-10070 (2005).
54. H. Masumoto, D. Hawke, R. Kobayashi, A. Verreault, A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* **436**, 294-298 (2005).
55. M. Giannattasio, F. Lazzaro, P. Plevani, M. Muzi-Falconi, The DNA damage checkpoint response requires histone H2B ubiquitination by Rad6-Bre1 and H3 methylation by Dot1. *J Biol Chem* **280**, 9879-9886 (2005).
56. I. Celic *et al.*, The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation. *Curr Biol* **16**, 1280-1289 (2006).
57. I. Celic, A. Verreault, J. D. Boeke, Histone H3 K56 hyperacetylation perturbs replisomes and causes DNA damage. *Genetics* **179**, 1769-1784 (2008).
58. H. N. Du, S. D. Briggs, A nucleosome surface formed by histone H4, H2A, and H3 residues is needed for proper histone H3 Lys36 methylation, histone acetylation, and repression of cryptic transcription. *J Biol Chem* **285**, 11704-11713 (2010).
59. S. Chen *et al.*, Structure-function studies of histone H3/H4 tetramer maintenance during transcription by chaperone Spt2. *Genes Dev* **29**, 1326-1340 (2015).
60. H. G. Tran, D. J. Steger, V. R. Iyer, A. D. Johnson, The chromo domain protein chd1p from budding yeast is an ATP-dependent chromatin-modifying factor. *EMBO J* **19**, 2323-2331 (2000).
61. Z. A. Chen *et al.*, Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.* **29**, 717-726 (2010).
62. S. Ovchinnikov, H. Kamisetty, D. Baker, Robust and accurate prediction of residue-residue interactions across protein interfaces using evolutionary information. *Elife* **3**, e02030 (2014).
63. Q. Cong, I. Anishchenko, S. Ovchinnikov, D. Baker, Protein interaction networks revealed by proteome coevolution. *Science* **365**, 185-189 (2019).
64. T. Gueudré, C. Baldassi, M. Zamparo, M. Weigt, A. Pagnani, Simultaneous identification of specifically interacting paralogs and interprotein contacts by direct coupling analysis. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 12186-12191 (2016).
65. S. Wang, S. Sun, Z. Li, R. Zhang, J. Xu, Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. *PLoS Comput Biol* **13**, e1005324 (2017).
66. H. Zeng *et al.*, ComplexContact: a web server for inter-protein contact prediction using deep learning. *Nucleic Acids Res.* **46**, W432-W437 (2018).

67. S. Wang, S. Sun, J. Xu, Analysis of deep learning methods for blind protein contact prediction in CASP12. *Proteins* **86 Suppl 1**, 67-77 (2018).
68. K. R. Roy *et al.*, Multiplexed precision genome editing with trackable genomic barcodes in yeast. *Nat Biotechnol* **36**, 512-520 (2018).
69. S. R. Collins *et al.*, Toward a comprehensive atlas of the physical interactome of Saccharomyces cerevisiae. *Mol Cell Proteomics* **6**, 439-450 (2007).
70. J. M. Schmiedel, B. Lehner, Determining protein structures using deep mutagenesis. *Nat. Genet.* **51**, 1177-1186 (2019).
71. N. J. Rollins *et al.*, Inferring protein 3D structure from deep mutation scans. *Nat. Genet.* **51**, 1170-1176 (2019).
72. R. W. Newberry, J. T. Leong, E. D. Chow, M. Kampmann, W. F. DeGrado, Deep mutational scanning reveals the structural basis for alpha-synuclein activity. *Nat Chem Biol* **16**, 653-659 (2020).
73. E. Eyal, R. Najmanovich, M. Edelman, V. Sobolev, Protein side-chain rearrangement in regions of point mutations. *Proteins* **50**, 272-282 (2003).
74. R. Sasidharan, C. Chothia, The selection of acceptable protein mutations. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 10080-10085 (2007).
75. J. Schaarschmidt, B. Monastyrskyy, A. Kryshtafovych, A. Bonvin, Assessment of contact predictions in CASP12: Co-evolution and deep learning coming of age. *Proteins* **86 Suppl 1**, 51-66 (2018).
76. S. Ovchinnikov, H. Park, D. E. Kim, F. DiMaio, D. Baker, Protein structure prediction using Rosetta in CASP12. *Proteins* **86 Suppl 1**, 113-121 (2018).
77. P. J. Robinson *et al.*, Molecular architecture of the yeast Mediator complex. *Elife* **4**, e08719 (2015).
78. J. Luo *et al.*, Architecture of the Human and Yeast General Transcription and DNA Repair Factor TFIIH. *Mol Cell* **59**, 794-806 (2015).
79. M. Jinek *et al.*, A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816-821 (2012).
80. J. P. Shen *et al.*, Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic interactions. *Nat. Methods* **14**, 573-576 (2017).
81. D. Du *et al.*, Genetic interaction mapping in mammalian cells using CRISPR interference. *Nat. Methods* **14**, 577-580 (2017).
82. L. Ma *et al.*, CRISPR-Cas9–mediated saturated mutagenesis screen predicts clinical drug resistance with improved accuracy. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 11751-11756 (2017).
83. A. V. Anzalone *et al.*, Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149-157 (2019).
84. D. E. Gordon *et al.*, A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459-468 (2020).
85. M. Eckhardt, J. F. Hultquist, R. M. Kaake, R. Huttenhain, N. J. Krogan, A systems approach to infectious disease. *Nature reviews. Genetics* **21**, 339-354 (2020).
86. D. E. Gordon *et al.*, Comparative Host-Coronavirus Protein Interaction Networks Reveal Pan-Viral Disease Mechanisms. *Science* **in press**, (2020).
87. D. E. Gordon *et al.*, A Quantitative Genetic Interaction Map of HIV Infection. *Mol Cell* **78**, 197-209 e197 (2020).
88. S. R. McGuffee, A. H. Elcock, Diffusion, crowding & protein stability in a dynamic molecular model of the bacterial cytoplasm. *PLoS Comput. Biol.* **6**, e1000694 (2010).
89. S. Takamori *et al.*, Molecular anatomy of a trafficking organelle. *Cell* **127**, 831-846 (2006).
90. B. G. Wilhelm *et al.*, Composition of isolated synaptic boutons reveals the amounts of vesicle trafficking proteins. *Science* **344**, 1023-1028 (2014).

91. J. Singla *et al.*, Opportunities and Challenges in Building a Spatiotemporal Multi-scale Model of the Human Pancreatic β Cell. *Cell* **173**, 11-19 (2018).
92. P. J. Thul *et al.*, A subcellular map of the human proteome. *Science* **356**, (2017).
93. M. Schuldiner, S. R. Collins, J. S. Weissman, N. J. Krogan, Quantitative genetic analysis in Saccharomyces cerevisiae using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. *Methods* **40**, 344-352 (2006).
94. W. Rieping, M. Habeck, M. Nilges, Inferential structure determination. *Science* **309**, 303-306 (2005).
95. W. Rieping, M. Habeck, M. Nilges, Modeling errors in NOE data with a log-normal distribution improves the quality of NMR structures. *J. Am. Chem. Soc.* **127**, 16026-16027 (2005).
96. H. Jeffreys, An invariant form for the prior probability in estimation problems. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **186**, 453-461 (1946).
97. A. Sali *et al.*, Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* **23**, 1156-1167 (2015).
98. S. K. Burley *et al.*, PDB-Dev: a Prototype System for Depositing Integrative/Hybrid Structural Models. *Structure* **25**, 1317-1318 (2017).
99. R. Chen, J. Mintseris, J. Janin, Z. Weng, A protein--protein docking benchmark. *Proteins: Struct. Funct. Bioinf.* **52**, 88-91 (2003).
100. C. M. Wood *et al.*, High-resolution structure of the native histone octamer. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **61**, 541-545 (2005).
101. A. Sali, T. L. Blundell, Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779-815 (1993).
102. S. M. Vos *et al.*, Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature* **560**, 607-612 (2018).
103. M. N. Wojtas, M. Mogni, O. Millet, S. D. Bell, N. G. A. Abrescia, Structural and functional analyses of the interaction of archaeal RNA polymerase with DNA. *Nucleic Acids Res.* **40**, 9941-9952 (2012).
104. D. Wang, D. A. Bushnell, K. D. Westover, C. D. Kaplan, R. D. Kornberg, Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941-954 (2006).
105. K. S. Murakami, X-ray crystal structure of Escherichia coli RNA polymerase σ70 holoenzyme. *J. Biol. Chem.* **288**, 9126-9134 (2013).
106. R. P. Joosten *et al.*, A series of PDB related databases for everyday needs. *Nucleic Acids Res.* **39**, D411-419 (2011).
107. W. Kabsch, C. Sander, Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **22**, 2577-2637 (1983).
108. M.-Y. Shen, A. Sali, Statistical potential for assessment and prediction of protein structures. *Protein Sci.* **15**, 2507-2524 (2006).
109. J. P. Erzberger *et al.*, Molecular architecture of the 40S· eIF1· eIF3 translation initiation complex. *Cell* **158**, 1123-1135 (2014).
110. Y. Shi *et al.*, Structural characterization by cross-linking reveals the detailed architecture of a coatomer-related heptameric module from the nuclear pore complex. *Mol. Cell. Proteomics* **13**, 2927-2943 (2014).
111. R. H. Swendsen, J. S. Wang, Replica Monte Carlo simulation of spin glasses. *Phys. Rev. Lett.* **57**, 2607-2609 (1986).
112. S. Viswanath, I. E. Chemmama, P. Cimermancic, A. Sali, Assessing Exhaustiveness of Stochastic Sampling for Integrative Modeling of Macromolecular Structures. *Biophys. J.* **113**, 2344-2353 (2017).

113. B. Vallat, B. Webb, J. D. Westbrook, A. Sali, H. M. Berman, Development of a Prototype System for Archiving Integrative/Hybrid Structure Models of Biological Macromolecules. *Structure* **26**, 894-904.e892 (2018).

114. J. D. Chodera, A Simple Method for Automated Equilibration Detection in Molecular Simulations. *J. Chem. Theory Comput.* **12**, 1799-1805 (2016).

115. L. McInnes, J. Healy, S. Astels, hdbscan: Hierarchical density based clustering. *The Journal of Open Source Software* **2**, 205 (2017).

116. E. D. Merkley *et al.*, Distance restraints from crosslinking mass spectrometry: mining a molecular dynamics simulation database to evaluate lysine-lysine distances. *Protein Sci.* **23**, 747-759 (2014).

117. D. Schneidman-Duhovny *et al.*, A method for integrative structure determination of protein-protein complexes. *Bioinformatics* **28**, 3282-3289 (2012).

118. M. J. de Hoon, S. Imoto, J. Nolan, S. Miyano, Open source clustering software. *Bioinformatics* **20**, 1453-1454 (2004).

119. A. J. Saldanha, Java Treeview--extensible visualization of microarray data. *Bioinformatics* **20**, 3246-3248 (2004).

120. J. H. Morris, C. C. Huang, P. C. Babbitt, T. E. Ferrin, structureViz: linking Cytoscape and UCSF Chimera. *Bioinformatics* **23**, 2345-2347 (2007).

121. N. T. Doncheva, K. Klein, F. S. Domingues, M. Albrecht, Analyzing and visualizing residue networks of protein structures. *Trends Biochem Sci* **36**, 179-182 (2011).

122. J. H. Morris *et al.*, setsApp for Cytoscape: Set operations for Cytoscape Nodes and Edges. *F1000Res* **3**, 149 (2014).

123. M. A. Collart, S. Oliviero, Preparation of yeast RNA. *Curr Protoc Mol Biol* **Chapter 13**, Unit13 12 (2001).

124. D. Kim *et al.*, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome biology* **14**, R36 (2013).

125. S. Anders, P. T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169 (2015).

126. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* **15**, 550 (2014).

127. M. Costanzo *et al.*, The genetic landscape of a cell. *Science* **327**, 425-431 (2010).

128. G. M. Wilmes *et al.*, A genetic interaction map of RNA-processing factors reveals links between Sem1/Dss1-containing complexes and mRNA export and splicing. *Mol Cell* **32**, 735-746 (2008).

129. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B* **57**, 289-300 (1995).

130. D. Wessel, U. I. Flugge, A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* **138**, 141-143 (1984).

131. M. Downey *et al.*, Acetylome profiling reveals overlap in the regulation of diverse processes by sirtuins, gcn5, and esa1. *Mol Cell Proteomics* **14**, 162-176 (2015).

132. J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* **26**, 1367-1372 (2008).

133. B. MacLean *et al.*, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966-968 (2010).

134. M. Choi *et al.*, MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524-2526 (2014).

135. R. Dronamraju *et al.*, Spt6 Association with RNA Polymerase II Directs mRNA Turnover During Transcription. *Mol Cell* **70**, 1054-1066 e1054 (2018).

136. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408 (2001).
137. R. Dronamraju, B. D. Strahl, A feed forward circuit comprising Spt6, Ctk1 and PAF regulates Pol II CTD phosphorylation and transcription elongation. *Nucleic Acids Res* **42**, 870-881 (2014).
138. J. T. Robinson *et al.*, Integrative genomics viewer. *Nat Biotechnol* **29**, 24-26 (2011).
139. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
140. R. Edgar, M. Domrachev, A. E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* **30**, 207-210 (2002).
141. J. Schneider, P. Bajwa, F. C. Johnson, S. R. Bhaumik, A. Shilatifard, Rtt109 is required for proper H3K56 acetylation: a chromatin mark associated with the elongating RNA polymerase II. *J Biol Chem* **281**, 37270-37274 (2006).
142. C. D. Kaplan, L. Laprade, F. Winston, Transcription elongation factors repress transcription initiation from cryptic sites. *Science* **301**, 1096-1099 (2003).
143. B. N. Devaiah *et al.*, BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin. *Nat Struct Mol Biol* **23**, 540-548 (2016).
144. M. S. Cosgrove, J. D. Boeke, C. Wolberger, Regulated nucleosome mobility and the histone code. *Nat Struct Mol Biol* **11**, 1037-1043 (2004).